LA-UR-17-25249
Approved for public release; distribution is unlimited.

Title: Next Generation Infrastructure Plan Level 2 Milestone Review

Author(s): Hick, Jason Cody
Vandenbusch, Tanya Marie

Intended for: Report

Issued: 2017-06-30

# Los Alamos
NATIONAL LABORATORY
EST.1943

Delivering science and technology
to protect our nation
and promote world stability

NNSA
National Nuclear Security Administration

Operated by Los Alamos National Security, LLC for the U.S. Department of Energy's NNSA

# Next Generation Infrastructure Plan Level 2 Milestone Review

## Advanced Simulation & Computing Program: Facility Operations and User Support


**Jason Hick, Tanya VandenBusch**

June 22, 2017

NNSA
National Nuclear Security Administration
Operated by Los Alamos National Security, LLC for the U.S. Department of Energy's NNSA

# Committee Members

- Randal Rheinheimer, LANL (Chair)
- Jim Lujan, LANL
- Paul Weber, LANL

Views expressed herein are personal evaluation expressed by the individual panel members and are not of any US Government entity or institutional entity.

# Acknowledgement and Recognition

- **Opening remarks (Jason Hick)**
  - Rationale
    - Impacting budgets (2 or more years out)
    - To balance the short duration projects with long-range projects
    - Improve coordination
  - Emphasis on building an understanding as well as communicating activities
    - Not just to produce a longer-term plan, but to create stronger synergies between program elements
  - Organization of paper/plan reflects current FOUS work packages
  - Series of meetings over the course of a month to understand and document how they plan, and what they propose for the next 2-5 years
  - Acknowledgement of efforts by FOUS managers & HPC leaders
    - Network - Susan Coulter, Jesse Martinez
    - Storage - Brett Hollander, David Bonnie, Kathy Sanchez
    - Operations - Andrew Montoya
    - Platform Environment - Cory Lueninghoner, David Montoya, Tim Randles
    - User Support - Ben Santos, Rob Cunningham
    - Monitoring - Mike Mason
    - Facilities - Phil Sena, Ron Velarde
- **Feedback on execution of review (Committee Members)**

# Assessment of Topics in the Agenda

- **Storage for ATS-1 and CTS-1 systems**
- **Networking to connect ASC infrastructure to LANL networks**
- **Facilities plans supporting future ASC systems**
- **Platform environment infrastructure for ASC systems**
- **Monitoring of infrastructure for improving HPC operations**
- **Operations plan for ASC systems**
- **Support for ASC users**

***To include:***

- *Have we identified the main drivers for each program element?*
- *What activities can help coordinate the plans of each program element?*
- *What are the challenges to consider with executing each program element's plan?*

# Storage for ATS-1 and CTS-1 systems

- **Home/Project**
  - Bandwidth is sufficient and don't want to encourage usage
  - Latency is important and can be improved by moving from Ethernet to IB
  - Capacity driven by ?, but is sufficient when enabling dedup/compression
  - Next 2-5 yrs focus enabling dedup/compression to satisfy capacity demands, and to work on replacing proprietary solution with homegrown to scale effectively in preparation for Crossroads
- **Scratch**
  - File systems purchased with machine acquisitions
  - Bandwidth determined with machine acquisition, based on system memory
  - Capacity driven by bandwidth
  - Plan is to support 10's of TB/sec for ATS-5 in 2020
  - Shared scratch (CTS systems) is determined by largest CTS system and is currently 100's of GB/sec
  - Next 2-5 yrs focus identifying suitable replacement in 4 years
- **Campaign**
  - Bandwidth goal will be order of magnitude less than scratch on largest system, so 1 TB/sec by 2020
  - Capacity goal is to store 1 year of data which represents 36 times (3 total system memories per month) the total system memory
- **Archive**
  - Focused on technology replacement to reduce cost
  - Changing usage from save everything to only what they need
  - Move away from interactive to batch usage

# Networking to connect ASC infrastructure to LANL networks

- **Campus network**
  - Ethernet based, 5-10 yr lifecycle
  - Driver is bandwidth between data storage & compute systems (data transfer)
  - Will not need replacing until FY19-22 timeframe
  - Next 2-5 yrs will replace turquoise firewall (FY18-19) with new high availability hardware solution
- **Storage network**
  - Can accommodate approximately 2 more additional compute clusters or file systems (depending on size)
  - Will not require replacement until FY20 timeframe
  - Next 2-5 yrs ongoing operations
- **External network**
  - DISCOM does not have requirement for increased bandwidth currently
  - Next 2-5 yrs DISCOM sites are qualifying 100Gb/s encryptors

# Facilities plans supporting future ASC systems

- **Facility projects**
  - Improving long-term value and capability of compute facilities
  - EC3E cooling capability, construction commences FY18 and needs to be complete by FY19
  - Initial planning for beyond exascale infrastructure requirements for FY25 (upgrade SCC, new building)
- **Installation projects**
  - Electrical connections for Crossroads in FY20
  - Mechanical/cooling connections for Crossroads in FY19
- **Next 2-5 yrs**
  - Redistribution of power within SCC
  - Project interface for above projects focused on HPC operations

# Platform environment infrastructure for ASC systems

- **Influenced heavily by platform deliveries**
  - Preparations for new environments
  - Involvement in Fast Forward, Design Forward, Path Forward, and other early system acquisition projects
  - Interacting with vendors and users for new tools and techniques
- **Participation in Tri-lab Common Computing Environment (CCE) efforts**
  - Common operating system
  - Scheduler
- **Exascale computing project is important to defining future software stack**
- **Next 2-5 years focus will be on**
  - Integrating new systems (CTS-2, ATS-3)
  - Understanding and improving user workflows
  - Improving speed of integration efforts, emphasizing testing
  - Completion of the Continuous Security Monitoring (CSM) effort to reduce time for security approvals
  - Improved accuracy and thoroughness during Dedicated System Times (DSTs)
  - Establishing a new partnership for SLURM scheduler enhancements

# Monitoring of infrastructure for improving HPC operations

- **System Monitoring**
  - 24/7 Monitoring and Log collection for all systems and infrastructure
    - Collaborate with CCE to establish Tri-Lab synergy for monitoring
  - All Monitoring systems and software were upgraded on March 2017
  - Next 5 years
    - Add Monitoring box per new cluster
    - 5 year infrastructure upgrade cycle
- **Facilities Monitoring**
  - Control and Monitoring Network for HPC Facilities
  - Next 1-2 Years require a complete redesign of network and infrastructure
- **Enhanced Monitoring**
  - Data Analytics Cluster
    - Used for Machine Learning and advanced analytics of HPC system data through collaboration with CSSE
    - Next 1-2 Years involve hardware purchase to support CSSE analytics
  - Continuous Security Monitoring
    - Upgrade of Security Plan through coordination with FOUS Platforms Environment
    - Next 1-2 Years involve completion of production deployment on all systems
  - Additional Data Sources
    - LDMS and Darshan prototype currently running on Woodchuck
    - Next 1-2 Years involve debugging and deploying into production on other clusters

# Operations plan for ASC systems

- **Incorporation of new tiered support**
  - Career and skills development for existing staff
  - HPC experience desired for new positions
  - Improve mean time to repair and responsiveness to many user tickets
- **Management of DSTs and outages**
  - Improving the state-of-practice for DSTs and outages to maximize work and minimize impact to users
- **Improving system availability**
  - Engagement in specific areas (file systems, archive) to reduce mean time to failure (swapping failed hardware components)

# Support for ASC users

- **Identify opportunities for improvement**
  - Communication during complex or protracted issues
  - Responsiveness, collaborating with new tiered support in HPC operations
  - Detailed problem diagnosis by coordinating distributed resources (AR team, consultant, and system administrators)
- **Identify the requirements for Crossroads**
  - Stay involved with AR team, NRE, and CoE efforts to understand Crossroads environment
  - New user environment, documentation, training?
- **Change storage usage**
  - Establish and communicate new policies for describing desired usage
  - Reinforce new usage

# Prioritized, Actionable Recommendations (3 to 7)

*(to achieve relevance to mission, leadership, efficiency, effectiveness, collaborations, etc )*

1. **FOUS should give more consideration to gathering, analyzing, and designing to, user and program requirements (top-down, e.g. consolidating machine rooms). The proposed infrastructure plan feels to be largely a continuation of existing capabilities (bottom-up), driven by the specific plan of record for future ASC platform acquisitions.**
2. **Production staff for platforms environment and platforms integration are often only engaged within six months of systems transitioning to production. It's not enough time. Having a plan to address platforms environment and integration sooner than 6-months before a system goes production would be a major benefit to production readiness for users.**
3. **It is clear from the plan that there is a new effort to move away from break-fix HPC operations to solving sysadmin level problems. It would be beneficial to see a more comprehensive multi-year Operations plan covering how to change to handling sysadmin level problems.**
4. **It seems that over the course of developing the plan, FOUS uncovered some deficiencies or idiosyncrasies related to its own internal structure and processes for planning and designing elements of the infrastructure. Suggestions for correcting some of these idiosyncrasies show up in various pieces of the plan. I would suggest that as a follow-up to this exercise, FOUS should initiate a more focused effort to assess the structure of the program, how the structure balances with recent changes to the structure of the HPC division, and what changes might be made to further improve communication and overall program efficiency. In concert, the infrastructure plan should remain focused on the future technical direction of the program.**
5. **HPC Division has run multiple iterations of similar strategic planning exercises, which have improved over time but could use further improvement. These are multi-programmatic and encompass a larger scope, but integrating the FOUS exercise into the HPC division strategic planning exercise would be helpful to both HPC division and the ASC Program. HPC Division also has a concise planning document indicating upcoming changes that contains exactly the kind of information captured in the details of the FOUS planning process. Integrating with and improving that reporting document would again usefully leverage existing processes.**

# Overall Comments or Minor Recommendations

1. FOUS has done a credible job of gathering future infrastructure requirements and identifying interrelationships and gaps, and has done so with a reasonable amount of time and effort. This effort is a good foundation for iterating to a recurring future planning process.
2. Several of the capability areas are on a technology refresh scale either closely or loosely tied to platform procurements, and there is clearly a set of periodic infrastructure enhancements that must all proceed together. Rather than organizing an infrastructure plan by capability areas, it might be useful to organize by platform cycle, identifying cycles that are similar to, or offset from, or independent of the platform cycle. This might naturally lead both to a plan more integrated among capability areas and one that indicates how budget smoothing could occur.
3. FOUS is the largest program impacting HPC operations within LANL. As such, strategic decisions by FOUS tend to heavily impact the strategic direction of all other local HPC programs. When planning future infrastructure decisions, it may be beneficial to integrate the requirements, goals, and concerns of the other programs. FOUS investments which can be made in concert or collaboration with other programs will naturally improve overall, lab-wide, and HPC division efficiency.
4. Consider how FOUS will identify user need for DISCOM bandwidth.
5. Better define infrastructure in the NGIP, especially in the HPC Operations section. It is used to describe both an organizational relationship and physical equipment. Ensure HPC division approves of new organizational changes.
6. Monitoring infrastructure seems mostly focused on the data gathering side. Infrastructure for analytics is also needed, and, more importantly for the overall monitoring system, infrastructure for returning insight from the analytics back to those who might use it to improve operations or a next procurement activity.
7. The “next 2-5 years” items in the platform environment element are either currently of interest or a current interest projected forward a few years. The only one that is not of that flavor is user workflows, which seems like it should be in the user support element.

# Thank you